



## King's Research Portal

DOI:

[10.1080/15248372.2015.1048864](https://doi.org/10.1080/15248372.2015.1048864)

*Document Version*

Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Burns, P., Russell, C., & Russell, J. (2016). Children's Binding of What-Where-When-Who in Episodic Memory: Identifying Oneself in the Past. *Journal of Cognition and Development*, 17(2), 300-319.  
<https://doi.org/10.1080/15248372.2015.1048864>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# Children's Binding of What-Where-When-Who in Episodic Memory: Identifying Oneself in the Past

Patrick Burns<sup>1</sup>, Charlotte Russell<sup>2</sup>, James Russell<sup>1</sup>

<sup>1</sup>Department of Psychology, University of Cambridge <sup>2</sup>Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College, London

Correspondence E-mail: jr111@cam.ac.uk.

## Abstract

It is usually accepted that the binding of What, Where, and When is a central component of young children's and animals' non-conceptual episodic abilities. We argue that additionally binding self-in-past (What-Where-When-Who) adds a crucial conceptual requirement, and ask when this becomes possible and what are its cognitive correlates. In the central task children between 3.5 and 6.5 years of age watched a light display on day-1, with two lights coming on simultaneously or in one of two orders. This was filmed from one of three positions: camera behind child, above child, and facing child. On day-2 children watched three videos from the original angle, each of which represented one of the three light configurations, and with the child in the video occluded. Participants had to decide which occluded child was them and justify their choice by reference to the lights. Above-chance performance was evident after 4.5- years. In addition, all children received the following tasks: spatial perspective-taking, seeing-leads-to-knowing, Modus Tollens reasoning, and second-order theory of mind. With age and verbal ability partialled out, only second-order theory of mind correlated significantly with performance on the central task.

Episodic memory is the conscious, re-experiential memory for autobiographical events. It can be viewed within cognitive development as something that depends upon the possession of certain concepts (such as theory of mind and causal insights into how experience causes memory: Perner, 2001), and as developing after 4 years (Perner & Ruffman, 1995). Alternatively it can be viewed as something more cognitively minimal and as essentially spatiotemporal (Russell & Hanna, 2012), with this form of episodic memory being found in children well under 4 years (Burns, Russell & Russell, 2014; Newcombe et al, 2014).

In fact, Endel Tulving's original definition of episodic memory was a minimalist, spatiotemporal one that made no explicit reference to conceptual abilities: "Episodic memory receives and stores information about temporally dated episodes or events, and temporal-spatial relations among these events" (Tulving, 1972, p. 385). Because of its minimalist nature, this definition has inspired research on forms of episodic memory in both animals and young children. While these two kinds of research differ in many respects, they hold the common assumption that minimal episodic memory is a form of "What-Where-When" memory. This paper focuses on what has to be added to children's What-Where-When memory through development so that it becomes a more conceptual and self-knowing kind of memory comparable to that of adults.

The What-Where-When shorthand in animal (Babb and Crystal, 2005; Clayton & Dickinson, 1998; Eacott and Norman, 2004; Iordanova, Good, and Honey, 2008) and developmental (Hayne & Imuta, 2011; Newcombe, Balcomb, Ferrara, Hansen, and

Koski, 2014) research means that episodic recall minimally depends upon binding spatiotemporal content (Where and When) to semantic content (What: an object or action). One immediate interpretive difficulty with this designation is that “temporal” can mean a number of different things. For example, does “temporal” mean how long ago the event took place (as in Clayton & Dickinson, 1998), the simultaneously present context of testing (as in Eacott and Norman, 2004; Newcombe et al, 2014), the day of the event (as in Iordanova et al., 2014), or the order in which locations were visited (as in Hayne & Imuta, 2011)?

In a recent deferred-imitation study with 2- and 3-year-old children, Burns, Russell, and Russell (2014) answered the question about the temporal element in the following way. “Temporal” should refer to the order or simultaneity of elements within the event, given that, as originally argued by Kant (1781/1998), the spatial content of the original experience will necessarily take this form. Using this criterion, we found evidence for a form of episodic WWW memory in 2- and 3-year-olds.

However, Burns et al used the term “proto-episodic” rather than “episodic” to designate the kind of memory displayed by the children. Reflection on the reasons for doing this will establish the rationale for the present study. In our experiment, the children had to recall, by re-enacting, the order (*When*) in which two levers were manipulated, the location of the relevant levers (*Where* on a box), and *What* kind of actions (pumping or twirling) that had to be performed. Overall, the children achieved this binding and their memory had an holistic character, with recall tending to be all-or-

nothing rather than fragmentary. There was, however, no reason to believe that they did all this *knowingly*.

That is to say, while the sight of the materials on the second day may have revived the experience of the experimenter's demonstration on the previous day (in perhaps a "flashback"-like way) there was no reason to believe that the children were demonstrating the following five kinds of conceptual ability that might be said to be the hallmarks of mature episodic memory:

- (1) Conceptualising the recollected event as a unique temporal locus in a sequence of events causally related to the recollected event (see Campbell, 1993; Hoerl, 2001; McCormack & Hoerl, 2001);
- (2) Intentionally casting their minds back to the previous day to report what happened, employing a cognitive agency that can be exercised additionally towards the future, sometimes called "mental time travel" (after Suddendorf & Corballis, 1997);
- (3) Construing their re-experiential memories in the following light: as the re-experience at time-2 of an original experience at time-1, thereby amounting to metarepresentation (Perner, 2001);
- (4) Regarding the experience at time-1 as *causing* the memory at time-2 (Perner, 2001);
- (5) Conceptualising *themselves* in the past as experiencers of the original event in a certain physical situation — "revisiting" themselves in the past as witnesses to the original demonstration.

Christoph Hoerl describes the significance of this last form of understanding in the following terms: “the causal understanding involved in episodic memory consists in a grasp of certain spatiotemporal constraints on remembering, that is, of the fact that *we must have been around to witness an event before we can remember it.*” (2001, p. 333; emphasis added).

The above five conceptualisations-cum-abilities are gathered together by Tulving (2005) under the term “autonoesis,” a coining that captures the necessity for the episodic recollector doing what she is doing with full self-knowledge. In this paper we focus on the “auto-” aspect — the conceptualisation of self-in-the-past — because this (the fifth above) conceptual ability would seem to underpin the other four. That is to say, thinking back to oneself as being in a certain situation yesterday would seem to subsume within it the following kinds of knowledge and ability: (1) regarding the past event as unique; (2) casting the mind back intentionally; (3) thinking of oneself as re-experiencing an earlier experience (metarepresentation); (4) regarding what one witnessed in the earlier situation as causing the present memory.

Given the generally-agreed centrality of spatiotemporal experience to episodic memory, as highlighted most recently in the Burns et al study, we decided to investigate the child’s ability to identify him or herself in the past (on a video record) by the use of spatiotemporal cues (simultaneity and order), which gives rise to the binding of the What-Where-When-Who of our title.

When do children manifest the kinds of conceptual abilities listed above, and what is the evidence that they are related to episodic memorial abilities? Taking free (versus cued) recall of coloured pictures as the measure of episodic recall Perner and Ruffman (1995) have shown that, when age and verbal ability are partialled out, there is a significant correlation with tests of understanding how earlier experiences determine later knowledge, which taps ability (4) above. When direct (versus indirect) exposure to test materials is the measure of episodic memory similar results are found (Perner, Kloo & Gornik, 2007). However, turning to ability (3) — metarepresentation — no evidence has emerged to date of a relation between first-order theory of mind as assessed by tasks such as unexpected transfer (Wimmer & Perner, 1983) and episodic recollection (Naito, 2003; Perner, Kloo & Stöttinger, 2007). The age of the children who are able to both perform well on the episodic tasks and pass the correlative cognitive tasks is 4 years or over.

The age of 4 years also emerged as a transitional age in a study by Scarf, Cross, Colombo, and Hayne (2013). In this, a “spoon test” (Tulving, 2005) methodology was employed as an index of the episodic recall of a single event: children had to select a key to open a locked box they had found previously. Because such a task required the exercise of the “cognitive agency” mentioned in ability (2) — casting the mind back to an event without perceptual reminders of it (in the service of an inference) — we would class it as a conceptual test of episodic memory, as opposed to the proto-episodic memory tapped in Burns et al (2014). The authors found, however, that when the delay between target event and item selection was short, and certainly under 24 hours, that many 3 year olds succeeded. This performance profile may have been due to the transfer

of the day-time event-memories from the hippocampus to neocortical regions during sleep (Burgess et al, 2001) in the older children. The younger children may have been failing to transfer the information overnight into long-term memory.

Age-of-acquisition issues also surround the ability to recognise oneself in the past from a video record which, as mentioned above, was a central requirement of our study. The seminal work on this topic has been carried out by Povinelli. Once again, the age of 4 is a transitional age. After 4 years children can appreciate the continuity between video-self and current self (Povinelli, Landau & Perilloux, 1996). Additionally, when children as young as 3 years are presented with videos of themselves in the past, they are able to self-recognise verbally (though often referring to themselves in the third person), whilst being challenged by the requirement to work out how the hiding event they saw in the self-including video determined the current state of the world (Povinelli et al, 1999).

### ***The Central Task***

In essence, the children (3.5-6.5 years) were required to decide, on day-2, whether they themselves or another child was present in a video of a visit to the lab on the previous day. The child in the video was occluded, and the only cues to identification that the video afforded were the spatiotemporal ones of whether a pair of lights the child in the video could see were coming on simultaneously, or left-before-right (LR), or right before left (RL). For example, if the child in the video was looking at simultaneous lights, and the child in the experiment could recall herself seeing the lights coming on successively, then the video could not be showing her.



We will refer to this task as “the jungle task” because, in order to engage the children, the lights were presented as “trees” in a “jungle.”

Given this procedure, it is clear that a child could fail the task simply because she had forgotten the light sequence she had seen. We felt, however, that explicit tests for sequence memory would fatally influence the measure of interest (use of spatiotemporal cues to self-identify), for the following reasons. First, giving a memory test before the test question would be sure to affect the answer given to the test question. Similarly, giving the memory test after the test question would be uninformative, as the child’s answer to the test question would constrain the response to the memory question. In addition, a memory check on day-1, after the initial demonstration, was not used as asking the self-identification question on day-2 could be contaminated by whatever memory the child had of the memory-check itself on day one. A virtue can be made out of this necessity, however, as it resulted in no cues being provided to encourage the use of the spatiotemporal information in self-identification, so we obtain a relatively pure measure of the spontaneous use of such cues —something more relevant to children’s everyday experience.

To avoid false positives from correct guessing we asked children to justify their answers; and these had to make reference to the light sequences. A further reason for requesting justifications was that we wished to tap children’s explicit rather than intuitive

understanding of how their choice was determined: autonoesis rather than simple memory.

We made the video recording on day-1 from one of three positions: behind the child, from the ceiling, and facing the child on the opposite side of the room. Given this, only in the first (behind) case could the video record provide a *familiar* image of the event. We will discuss the role of camera positions later; but *prima facie* the camera-behind condition would seem, at least, to be easier given that the child is being presented with a scene that should look familiar, as it is an image of the original experience.

To pass the task then, children had to identify the video that contained the arrangement of lights to which they had been exposed, in addition to justifying their choice with reference to the lights.

### ***Ancillary Tasks***

In addition to determining when children can spontaneously use past spatiotemporal cues to self-identify we wished to find out what were the cognitive correlates of this ability.

To this end we gave children four further tasks, in addition to a test of verbal ability. We now present the tasks and their rationales.

### ***Modus Tollens***

Making a correct video choice in the self-identification task requires the following kind of inference: “If it is me in the video then there will be left-right lights shown, there were

simultaneous lights, so this cannot be me.” This inference takes the form *If p then q, ~q, therefore ~p*, which is Modus Tollens.

### ***Coding Spatial Experience In Terms Of One’s Location***

Success on the task would also seem to depend on the child’s ability to reason about the location of objects in relation to an observer at different locations. We adapted a task used by Newcombe and Huttenlocher (1992) to this end.

### ***Second-Order Metarepresentation***

As mentioned above, no evidence has emerged of to date for a linkage between episodic memory development and performance on classical metarepresentation tasks like the unexpected transfer task for false belief understanding (Naito, 2003; Perner et al, 2007). But in fact success on the self-identification task would seem to require not only metarepresentational ability (as in representing at time-2 what was experienced at time-1; e.g., simultaneous lights) but *second-order* metarepresentational ability, namely, representing a representation at time-2 of one’s mental representation of simultaneous lights at time-1). Accordingly, we presented children with a task based on Perner and Howes (1992) in which children had to embed one propositional attitude within another.

### ***Seeing Leads To Knowing: Causal Knowledge***

Recall that one of the five kinds of conceptual appreciation discussed above (number 4) was one involving the knowledge that the experience at time-1 was the cause of the re-experiential memory at time-2. In tasks assessing this children have to say whether they

knew or guessed the location of a toy. Perner and Ruffman (1995; though see Drummy & Newcombe, 2002, for somewhat conflicting data) found a correlation between episodic memory (assessed by free recall for pictures) and understanding that knowing something at time-2 depends upon have experienced (told or seen) something at time-1. We presented children with these “see-know” tasks based on Perner’s procedures.

Our two central questions for this study were then: (1) When do children become capable of passing an episodic recollection task in which there must be explicit representation of self-in-past? (2) What kinds of conceptual abilities seem to be underpinning this ability? We tested children between 3.5 and 6.5 years.

## METHOD

### *Participants*

We tested 166 children (85 female) between the ages of 3.5 and 6.5 years of age. Thirteen children were removed from the final sample either because a parent or an older sibling interfered with experimental protocol or because the child refused to participate in one of the tasks. The final sample consisted of 153 children, 51 between the age 3.5 and 4.5 years ( $M = 48$  months), 51 between the age 4.5 and 5.5 years ( $M = 58$  months) and 51 between the age 5.5 and 6.5 years ( $M = 74$  months). These age-bands will be referred to henceforth as “young,” “middle,” and “old.” A third of the children ( $n = 17$ ) in each age-band were randomly assigned to one of the three conditions. Children were recruited via leafleting and posters in schools and children’s centres in Cambridge, UK. The SES of the children was predominantly middle class.

### ***Apparatus***

A playroom in the laboratory was employed as the “jungle room.” The room was decorated with potted plants and soft toys to give the impression of a jungle (see Figures 1 and 2). In the centre of the room there were two identical light towers (schematic “trees”). They were cuboid in shape and measured 1ft x 1ft around and 4ft in height. The base of each tower was painted green while the top part consisted of an opaque plastic cover underneath which there was a red light. There was a small chair in the room for children to sit on and two larger chairs against a side wall on which parents could observe. In the centre of the room, 2.5 feet apart, stood the two light towers. The chair the child sat in was placed 5ft in front of the light towers. In the camera-behind and camera-facing conditions a Panasonic camcorder was used to record the child’s visit. The camcorder was positioned on a tripod either directly behind the child’s seat in the *camera-behind* condition, or facing the child on the far side of the room, 5ft from the light towers, in the *camera-facing* condition. The height of the camcorder was fixed at just above the child’s seated head-height. For the *camera-above* condition the ceiling-mounted camera was above the towers and recorded an image of the whole room.

### ***Design And General Procedure***

In addition to age, the between-subject factors were (a) camera position and (b) the temporal order of the lights (simultaneous, left-before-right, or right-before left).

See Figure 1. With regards to (a), in the camera-behind condition, the camera recorded the episode from the same perspective as that of the child. In the camera-above condition a ceiling camera recorded the episode from above (90° rotation upward from the child's perspective) and in the camera-facing condition a camera recorded the scene from the opposite side of the room from where the child was positioned (180° rotation from the child's perspective). Note that in the camera-above condition there was congruence between the child's left-right experiences and the camera record, despite the fact that the views were different. Accordingly, what was, say, left to the child in the jungle visit was left on the camera-above clip.

The two light towers ("trees") were placed in the centre of the room, surrounded by plants and stuffed animals. (The latter changed position from child to child, depending on what happened in the free-play session — see below.) The child sat facing the towers, equidistant from them such that one tower was in the child's left field of vision and one tower was in the child's right field of vision (see Figure 1). The focal event was the light display that accompanied the monkey calls. Children either saw the left light come on for 5 seconds after which there was a pause of one second before the other light came on for 5 seconds. Alternatively, they saw the right light come on first followed by the left light. Or, thirdly, they saw both lights come on simultaneously for 5 seconds.

On day-2 children were shown 3 video clips of the jungle visit, each one taken from the camera position of the experimental condition to which the child had been assigned. A piece of blank card was placed over the screen, occluding the child present in

the video. In fact, there was no child present in any of the three video clips as they were shot immediately after children's visit on day 1 was completed. The 3 clips were identical except for the spatiotemporal content of the light sequences. The video clips were played without sound. Children selected the video in which they thought that they were present. Having selected one of the three videos in response to the choice question, children were then asked why they had made that choice. Responses to the justification question were categorised as either a correct or incorrect justification. Children were credited with giving a correct justification if they made reference to the temporal order of the lights. Examples of correct justifications were, "Because that one went on first" (pointing to one of the lights), and "Because the two lights came on at the same time". Incorrect justifications made no reference to the temporal sequence of the lights: for example, "Just know", "It had the monkeys in it" and "Because the lights were that colour". Including a justification of the choice limited the influence that picking the correct video by chance<sup>1</sup> could have on the data.

Children completed 5 further tasks (described below): second-order mental state attribution (Perner & Howes, 1992); visual perspective taking (Newcombe & Huttenlocher, 1992); know-guess (Perner & Ruffman, 1995); Modus Tollens reasoning (modelled on McCormack, et al, 2013); British Picture Vocabulary Scale 1<sup>st</sup> edition (BPVS). Two of the tasks were administered on day-1 after children had visited the jungle and the remaining three tasks were administered on day-2 after the children had

---

<sup>1</sup> For example, if 50% of the children selected the wrong video and if we assume that these children chose randomly from the three options, with a chance of 0.66 of being wrong then we can estimate that approximately 25% of the whole group selected the correct video by chance (given that the chance of a correct guess, 0.33, half of 0.66).

watched the video clips of the jungle visit. The tasks were administered in a randomized order with the exception of the BPVS, which was always administered last.

### ***Narrative Procedure For The Central And Ancillary Tasks***

When children arrived they were told that they were going to visit a pretend jungle where they would play a game in which they had to identify various animals from the calls they made. They were told about a special animal noise that they had to listen out for: that of the monkey. The experimenter also told children that when the monkey calls were heard “something special will happen,” though he “wasn’t sure what kind of special thing would happen.” The experimenter then brought children accompanied by their parent/caregiver to the jungle room. After some warm-up play with the toy animals the main game began. The experimenter first introduced the camera to children, telling them “We will make a video of our trip to the jungle so that tomorrow we can watch the video and see ourselves playing in the jungle.” The experimenter then turned on the video. In the camera-behind and camera-facing conditions the camera was placed on a tripod and the child was invited to look through the viewfinder and asked if he or she could see the jungle trees and the experimenter (who positioned himself in front of the camera). In the camera-above condition the child was brought into an observation room adjacent to the lab (behind a wall to the right of the child’s seat in Figure 1) where the video equipment for the ceiling camera was located. The experimenter pointed out on the monitor the jungle trees and the small chair where the child would be seated. The experimenter then went back into the jungle and waved into the camera (as in the camera-facing and



camera-behind condition) and asked the child to verify that he or she could see him on the monitor.

The child then returned to the seat in the jungle to begin the game. Two recordings of animal noises were played: first an elephant and then a parrot. Children were given clues if they failed to identify either animal. The third animal noise played was that of the monkey. Children were asked prior to the monkey call if they could remember which animal was the special one that they had to listen out for, and were given a reminder if they could not remember. The monkey recording played in exact synchrony with the lighting of the “tree” towers. Once the display had finished the experimenter asked the child whether he or she had seen the lights and commented on how surprising it was. The experimenter then told the child that he thought that was all the animals they would hear. He switched off the camera while reminding children that they would be able to watch the video tomorrow. The experimenter, child and caregiver then moved to a separate room where they completed two other tests.

Children and caregivers returned approximately 24hrs later. The experimenter reminded children that the day before they had visited the jungle and that the experimenter had made a video. The experimenter then told the child that he was going to show them a video clip from their visit to the jungle the previous day together with two clips from the visits of other children. They had to pick the video clip in which they were present. The videos were presented on a 14-inch laptop screen. A card was placed in front of the portion of screen where a child would be. The order in which the three videos were

shown to children was counterbalanced. In each condition, before the first video was played the experimenter pointed out the jungle trees and the 'jungle lights' to children. Children were also told that someone was sitting in a little chair behind the card and that it might be them but it might be someone else. The experimenter then played the video clip, which began just before the monkey call commenced and finished just after the monkey call finished. When all three clips had been viewed the experimenter asked children, "Which one of these three video clips is our video?" If children gave no response they were asked whether they wanted to see the videos again. They were then shown the 3 video clips in the same order with the same instructions, after which the test question was put to them again. If they did not reply to the test question after further encouragement, they were recorded as giving no response. When children selected a video clip the experimenter replayed the clip, asked children whether they thought this was the right one and then asked children *why* they thought this was the right video. Children were then shown the entire film of their own visit to the jungle. Children then completed the remaining tasks. The procedures for each of the 5 ancillary tasks are described below.

*Second-order mental state attribution* (after Perner & Howes, 1992): Children were told a story, illustrated with pictures, about a boy who received a new toy train from his mother. After playing with his train, he left it on the table in his bedroom and then went to the park to play with his friend. While he was in the park his mother came into to his room to tidy up. She saw the train on the table and she moved it to a box on the shelves. Children were asked two control questions: "Where did Charlie [the boy] leave his train?" and

“Where is his train now?” They were then asked the second-order mental state attribution question in which giving the correct answer requires the child to appreciate that the protagonist *thinks* he *knows* where the train is. They were told to imagine going to the park and meeting Charlie so they could ask him: “Charlie, do you know where your train is? What will Charlie say? Will he say, ‘I do know where my train is or will he say ‘I don’t know where my train is?’” Once they had made a response children were then asked why they thought Charlie would say that. Answers to the why-question were coded as appropriate if they made reference to Charlie’s mental state, “He thinks it is on the table”, or if they made reference to Charlie’s leaving the toy at a particular location, “Because he left his train on the table”. Inappropriate justifications were those that referred to the toy’s current location, “Because the toy is in the box”. As a final check, children were asked a first-order false belief question, “Where will Charlie look for his train when he goes back home?” Children who gave an inappropriate justification and/or answered the first-order false belief question incorrectly were categorised as failing the task. Those who answered the second-order mental state attribution question correctly, gave an appropriate justification and answered the first-order false belief question correctly were categorised as passing the task.

*Visual perspective taking* (after Newcombe & Huttenlocher, 1992): Children sat at one side of a small square table (notional North) in a yellow chair. The experimenter sat at the notional West side of the table. On the South side directly opposite the child there was a red chair. The experimenter and child each put on a single glove (the child chose left or right with the experimenter copying the child’s choice). The gloves were used because

young children struggle with the terms “left” and “right.” The experimenter then introduced four coloured blocks, placing one at each side of the table. Children placed their hands on the table in front of the East and West blocks. The experimenter then labelled the blocks’ location for the child: “The yellow block is right in front of you. The red block is far away from you. The blue block is by your hand with a glove and the green block is by your hand without a glove.” Children were then asked four control questions, such as “Which block is right in front of you?” and so forth. If they answered any incorrectly they were then reminded of each block’s location and then the four control questions were repeated. Before the test questions were asked they were told to pretend that they were sitting in the red chair. They were then asked four test questions presented in conditional form, “If you were sitting in the red chair, which block would be right in front of you / far away from you / by your hand with a glove / by your hand without a glove?” The experimenter held his glove-hand or non-glove-hand up in front of his face while asking those questions. Children were given a score out of four.

*Know-guess task* (after Perner & Ruffman, 1995): Two boxes, one red and one blue, a toy horse and a screen were used. The experimenter told the child that he would put the horse in one of the 2 boxes and that their task was to then find it. There were three types of trial: (1) a ‘see-know’ trial in which the boxes were positioned in front of the screen in full view of the child such that the child could see in which box the horse has been placed; (2) a ‘hear-know’ trial in which the boxes are positioned behind the screen but the experimenter told the child in which box the horse has been placed; (3) an ‘ignorance’ trial in which the boxes were positioned behind the screen but the child was

not told in which box the horse had been placed. Each trial was completed twice. The ignorance trials were rigged such that on one the child was sure to select a box containing the horse (both boxes contained a horse) and on one the child was sure to select the box without a horse (neither contained a horse). Children were asked which box contained the horse. Once they had selected a box and it was opened they are they were asked the test question: "When I asked where the horse was, did you know the horse was in this box or did you guess?" On the ignorance-correct trial, children who claimed they knew the location of the horse were asked a follow-up question: "How did you know the horse was in the red/blue box?" Responses were scored in the following manner. Children were given a score of 2 for correctly asserting that they knew where the horse was on the 'know' trials and guessed which box the horse was placed in on the ignorance trials. Children were given a score of 1 for claiming they had known where the horse was on the ignorance-correct trial while appropriately selecting guess on the ignorance-incorrect trials and know for the remaining know-trials. However, if on follow-up questioning, children gave a plausible reason for claiming that they knew where the horse was on the ignorance-correct trial then they were given a score of 1.5. (Plausible reasons included inferring the location of the horse from the pattern of previous hiding places, e.g., "I knew it was going to be red this time because you put it was in blue the last two times", and by careful observation, "I saw your hand move to that one"). Finally, children were given a score of 0 for responding 'know' or 'guess' to all test questions or for selecting know and guess without any discernible pattern.

*Modus Tollens* (modelled on McCormack et al., 2013): Children were presented with two tests of Modus Tollens reasoning. On the “card” test children were shown an envelope and told that it contained some cards. The experimenter told children that some of the cards were red and some were blue (withdrawing one card of each colour to demonstrate and then placing them to one side). The experimenter then told children that there was a special rule with the blue cards: “If the card is blue then there is a picture of a teddy bear on the other side”. The experimenter then withdrew a blue card and asked the child what is on the other side. They were reminded of the special rule if they did not give the correct response. When they gave the correct response the card was turned over to reveal the picture. The experimenter repeated this with another blue card and then on the critical trial the experimenter withdrew a card with a picture of a *cat* ( $\sim q$ ) on it. The experimenter then asked the test question: “Is this a blue card or is this a red card?” The “box” test was identical except that it involved black boxes that were either yellow or green underneath. Green boxes contained marbles. On the critical trial a box is opened to reveal a *crayon* ( $\sim q$ ) and children were asked whether it is a green or yellow box. Scoring ranged from 0-2.

## RESULTS

### *Performance On The Jungle Task*

In analysing performance on the jungle task we used two outcome measures: (1) children’s choice of video and (2) video choice plus children’s justifications of their choices (a selection of these justification is shown in an appendix). Table 1 presents the percentages of children, across age-group and camera position, selecting the correct

video. In parenthesis are the percentages of children selecting the correct video and also giving an appropriate justification for their choice. Table 2 presents the percentage of correct video selections (and video selections with appropriate justifications in parenthesis) across age group and light display order. We did not combine the data from Tables 1 and 2 into a single analysis as this would have resulted in cell sizes too small to make for meaningful analysis.

We first examined the effect that age-group and camera position had on correct video choice, without taking justifications into account. We conducted a loglinear analysis to examine the 3-way interaction of age-group (young vs. middle vs. old), camera position condition and outcome (pass/fail) and all lower order effects and interactions (age-group x position; age-group x outcome; position x outcome; age-group; position; outcome). The final model only retained the effect of age-group by outcome. The likelihood ratio of this model was  $\chi^2(12) = 4.59, p = .97$ , indicating that this model was a good fit for the data. The 2-way interaction term of camera position by outcome was not retained, indicating that camera viewing position, unlike age, had no reliable effect on correct video choice.

To explore this effect of age on performance we conducted three chi square analyses.

There was a significant difference in the proportion of young and middle age-group children who selected the correct video,  $\chi^2(1) = 7.02, p < .01$ . There was also a significant difference in the proportion of young and oldest age-group children who selected the correct video,  $\chi^2(1) = 13.02, p < .01$ . There was no significant difference in

the proportion of middle and oldest age-group who selected the correct video,  $\chi^2(1) = 1.06, p = .3$ .

We next examined the effect that age and camera position had on correct video choices *with* justifications. A loglinear analysis examined the 3-way interaction of age-group, camera condition and correct response with justification, along with each of the lower-order interactions and effects. The final model retained the effect of age by correct response-with-justification. The likelihood ratio of this model was  $\chi^2(12) = 7.72, p = .81$ , indicating that this model was a good fit for the data. That the 2-way interaction term of condition by correct response with justification was not retained indicates that camera position had no effect on this measure of performance.

The interaction of age by correct response-with-justification was explored with 3 chi-square analyses. The proportion of middle age-group children who gave a correct response with a justification was greater than the proportion of younger age group children,  $\chi^2(1) = 5.45, p < .05$ . Likewise, the proportion of older age-group children who gave a correct response with a justification was significantly greater than the proportion of middle age-group children,  $\chi^2(1) = 3.97, p < .05$ . Lastly, the proportion of older age-group children giving correct responses with appropriate justification was significantly greater than the proportion of younger age-group children doing so,  $\chi^2(1) = 17.49, p < .01$ .



We next examined the effect that age and *light display type* had on correct video choice, discounting justifications. A loglinear analysis examined the 3-way interaction of age-group by light display type (left-right/simultaneous/right-left) by outcome (pass/fail) and each of the lower type interactions and effects. The final model retained the interaction of age by outcome only. The likelihood ratio of this model was  $\chi^2(12) = 14.02, p = .3$ , indicating that this model was a good fit for the data. Critically, the interaction of light-display type by outcome was not retained, indicating there was no association between the two variables. The interaction of age with correct video choice has already been explored above.

Finally, with regard to jungle task performance, we examined the effect of age and light-display-type had on our second outcome measure: correct responses *with* appropriate justifications. A loglinear analysis examined the 3-way interaction of age by light display type by correct response with justification and each of the lower order interactions and effects. The final model retained the interaction of age-group by outcome only. The likelihood ratio of this model was  $\chi^2(12) = 11.17, p = .52$ . The 2-way interaction term of light-display-type by correct response with justification was not significant indicating, that there was no association between the kind of light display witnessed and success on the jungle task.

### ***Possible Reasons For Failure: Not Encoding The Lights And Reversing Onset Orders***

We considered two *prima facie* reasons why children should fail the task: failure to notice the lights coming on, and confusions of left and right. We give descriptive statistics bearing on these concerns.

In the first place, 19% of children made spontaneous reference to the lights coming on, and 31% spontaneously pointed to the lights. The large majority of children (86%) reported that they had seen them come on when asked. The ones who did not reply tended to be generally uncommunicative children. One child said he did not see the lights come on, so the sequence was replayed for him.

Thirty-four percent of all responses were errors on the successive conditions (LR or RL). See Table 3 for these errors against age. These differences in error type against age fell short of significance:  $\chi^2(2) = 5.44, p = .066$ . Of these errors, 64% were errors of choosing the video with the *simultaneous* lights, rather than reversal errors. The remainder (36%) were indeed reversal errors. Given this pattern, it can be said that confusing left and right was not the major source of error even in the youngest children. Indeed, in the camera-facing condition (where the observed light order was the *opposite* from the one originally seen) there was the smallest number of reversal errors: 11%, with 44.5% each on the other two camera positions.

### ***Video Selection Against Chance***

A small percentage of children (7%) said that they did not know when asked to select one of the videos. Excluding those, we can examine whether the remaining children in each

age group selected the correct video more often than chance (chance is 1/3). Children in the young group were no better than chance at selecting the correct video (one-tailed binomial test,  $p = .22$ ). Middle-group children (3.5-4.5) were, however, significantly better than chance at selecting the correct video (one-tailed binomial test,  $p < .01$ ) as were children in the oldest group (one-tailed binomial test,  $p < .01$ ).

### ***Performance On The Ancillary Tasks***

Table 4 provides a summary of children's performance on the four ancillary tasks of second-order theory of mind, Modus Tollens reasoning, perspective-taking and see-know/guess. These data were used in the analysis. It is notable that correct responding on the Modus Tollens task was near ceiling level for even the youngest age group. (The perspective-taking and know-guess data are presented the way they are because the data were strongly bi-modal.)

### ***The Relation Between Jungle Task Performance And Ancillary Task Performance***

We examined the relationship between children's performance on the jungle task and the ancillary tasks. We used children's choice of video plus appropriate justifications as the performance measure in the jungle task as this is a conservative measure of performance. Table 5 presents full and partial correlations between the different measures after controlling for age and receptive vocabulary. This demonstrates that, when age and verbal ability were partialled out, only performance on the second-order theory of mind task correlated significantly with performance on the jungle task (column 3, lower

portion). The other correlations fell well short of significance. To examine this result more closely we carried out a logistic regression analysis.

We performed a hierarchical mixed binary logistic regression with performance on the jungle task as the dependent measure. A regression analysis models the influence of multiple variables on the dependent measure and allows us to predict the probability of new participants passing the jungle task. In step-one, age (measured in months) and BPVS score were entered into the model. The likelihood ratio test confirmed that the model at this stage was a significantly better predictor of the recall in the jungle task than the null model,  $\chi^2(2) = 32.94, p < .01$ . (The null model is one in which no predictor variables are included and each participant is assigned category membership based on the modal outcome, i.e., failing the task). The Wald chi-square statistic indicated that BPVS was a significant predictor of the outcome,  $\chi^2(1) = 10.85, p < .01$ . Age alone was, however, not a significant predictor of the outcome  $\chi^2(1) = 2.61, p = .11$ . In other words, when predicting performance on the jungle task it is no more beneficial to know age if verbal ability is known: the effect of age is fully explained in terms of an increase in verbal ability. In step two we entered theory of mind, know/guess, perspective taking and Modus Tollens reasoning into the model using the backward stepwise likelihood ratio method, whereby all variables are entered into the model to begin with and then each variable is removed in turn. The variable least likely to predict the outcome was removed at each iteration and the model without that variable was compared to the model with that variable. If there was no significant difference in the variance predicted by the two models then the removed variable is not retained. At each iteration, then, we determined

how well the model fitted the observed data. Only second-order theory of mind was retained in the final model. The likelihood ratio test confirmed that this model was a significant improvement on the model at stage one (when only BPVS and age were included):  $\chi^2(1) = 4.81, p < .05$ . The final regression model is presented in Table 6. Compared to the null model the final model was a significantly better predictor of performance in the jungle task:  $\chi^2(3) = 38.39, p < .01$ . That is, it accounted for a significantly greater proportion of the variance on the jungle task than did the null model. The Hosmer and Lemeshow test indicated that the final model provided a good overall fit for the data:  $\chi^2(8) = 7.99, p = .43$ . In other words, the predicted values of each case derived from the model were not significantly different from the observed values. The Nagelkerke R Square value associated with this model was .311. This is an analogue of the  $R^2$  value in linear regression, indicating that almost a third of the variance in performance on the jungle game is accounted for by the model. Analysis of the individual predictor variables in Table 4 indicates that performance on the BPVS and theory-of-mind task were the only significant predictors of recall in the jungle task. The positive regression coefficient  $\beta$  and an odds ratio of greater than 1 indicates that increases in BPVS and theory of mind are associated with increased probability of recall on the jungle task; i.e., as scores on BPVS and second-order theory of mind increase so too does performance on the jungle task.

## DISCUSSION

Two clear results have emerged from this study. First, children above 4-and-a-half are likely to succeed on a task in which they have to infer their presence in a past

spatiotemporal context from video evidence. Second, while a grasp of the see-know relationship, skill at spatial perspective-taking, and Modus Tollens reasoning were not related to performance on the jungle memory task, second-order theory-of-mind performance was indeed related. Less centrally, we have broadly replicated the results of earlier studies on Modus Tollens understanding, spatial perspective-taking, seeing-leads-to-knowing, and second-order theory of mind in children of within this age range.

### ***The Significance Of Success After Four-And-A-Half Years***

In the first place, the age at which performance on video selection rose above chance in the jungle task is not markedly different from the age at which children succeed on other “conceptual” (see Introduction) tests of episodic memory using free recall (Perner and Ruffman, 1995), direct-versus-indirect experience (Perner et al, 2007), and the spoon-test methodology (Scarf et al, 2013). It is likely, then, that requiring children explicitly to represent themselves in the past in an episodic memory task makes no significant cognitive demands additional to those recruited by these other conceptual episodic tasks. This is consistent with the view that this is something that they do *naturally* when presented with the other episodic tasks. Second, research by Povinelli, alluded to in the Introduction, tends to suggest that the children in our sample were unlikely to have been impeded by an inability to relate the self-in-video to present, remembering-self. Children of 4 can appreciate the continuity between video-self and current self (Povinelli, Landau & Perilloux, 1996). While even children of 3 years show verbal recognition of themselves in videos (often referring to themselves in the third person); though it will be

another year or two before they can understand the causal linkage between the past, videoed configuration of the world and the current one (Povinelli et al, 1999).

### ***The Role Of Second-Order Tom***

Can one conclude from these data that second-order theory of mind (the embedding of “know” within “think” in this case) is implicated in the child’s developing ability explicitly to place herself in a past spatiotemporal context as an experiencer? It might be objected that we found the result we did because the task was *designed* to require second-order cognition, at least in the camera-above and camera-facing conditions, implying that such a relationship is not surprising and that it has no useful implications for the normative development of episodic memory. There are two replies to this point. In the first place, to find a significant correlation (after age and verbal ability have been controlled for) between a task that requires the embedding of one mental orientation within another and one that requires working out the identity of a filmed agent from spatiotemporal cues is hardly trivial. The tasks are markedly different. Second, and more centrally, the camera-above and the camera-facing conditions were necessary inclusions, as we have indicated, because in the camera-behind condition, when child and camera shared a perspective, it was possible for the child to be correct simply by virtue of the video showing a *familiar* scene. It was not necessary for the child to have any grasp of the fact that her location in a past spatiotemporal context determined a certain kind of experience.

However, if the camera-above and camera-facing conditions required a more demanding strategy from the children, why did we not find a significant difference in performance between the three camera-placement conditions (see Table 1)? Why was the camera-behind condition not easier than the other two conditions? A strong possibility is that the successful children were *not*, in fact, performing in the camera-behind condition on the basis of familiarity alone, but by a form of second-order cognition of the kind “Am I seeing the view representing the representation of the lights that I had yesterday.” In other words, the result demonstrates the emergence of a mature form of episodic memory.

### ***Perspective-Taking In Performance On The ‘Camera-Facing’ And ‘Camera-Above’ Conditions***

The discussion in the previous section carries the assumption that in the camera-above, and especially in the camera-facing conditions, successful children engaged in a kind of transformation of egocentric perspectives. It must be born in mind, however, that the room also contained *allocentric* cues such as windows, doors, plants, and toy animals. Accordingly, a child in conditions other than the camera-behind one could plausibly represent something like “the first tree lighting up was next to the flamingo” (see Figure 2). This possibility does not, however, vitiate the claim that in the camera-above and camera-behind conditions the successful children could have been engaging in second-order representation. For example, such a child would still have to do the following in the (say) camera-facing condition: “I am seeing/representing the flamingo-light-first



representation (1), which would also have been represented (2) by the hidden child.”

This role of this “transformation of egocentric perspectives” will be discussed again later.

### ***The Relation To The Metarepresentational Theory Of Episodic Memory***

How do these data relate to Perner’s (2001) claim that conceptual episodic memory depends upon what is essentially a first- (not second-) order understanding? Essentially, on Perner’s view, two forms of understanding have to be in place for episodic memory:

(1) understanding what it means for something to be representation of a state of affairs (this may not be mental, and may be pictorial: Perner, 1991), (2) understanding that experiences at time-1 can cause memories at time-2. On our view, however, there is an additional representational step, which is representing the representational states of oneself at time-1.

It may be objected to the current position that perhaps our study, in virtue of the additional demands of the camera-above and the camera-facing condition, added a representation step that need not be present in the normal case. Indeed a sceptic might say that the whole notion of binding “past-self” to past spatiotemporal contexts is somewhat overblown? After all, to whom would the child bind these experiences if not to herself?

To answer this criticism one must take account of the fact that episodic memory is (a) from a point of view and (b) includes the bodily location of the subject. These two things will be true of it if indeed the rememberer is including herself in the representation as a physically-present witness. The first-order account omits (a) and (b) insofar as it is possible for one to take oneself to be enjoying a re-experiential memory of a scene

*without* the added appreciation that (say) X was experienced on the left because of one's spatial location in relation to X and other objects. To quote Hoerl again (2001, p. 333) "the causal understanding involved in episodic memory consists in a grasp of certain spatiotemporal constraints on remembering, that is, of the fact that we must have been around to witness an event before we can remember it... This kind of causal understanding may be quite distinct from the ability to think of the causal relations one's mental states stand to each other." Given this, it is not enough to refer to mental-state-at-time-one causing mental-state-at-time-two. One has to refer also to the past location of the experiencing body, something that was required, albeit not consistently, in the jungle task.

***The "Transformation Of Egocentric Perspectives:" Neuroscientific Implications.***

The phrase "transformation of egocentric perspectives" was used in the section before last to describe what children were being called upon to do, most clearly in the camera-above and the camera-facing conditions. That is to say, in these two conditions the child not only had to operate with an allocentric representation of the display (e.g., child-who-may-be-me between the two towers) but had mentally to compare this to their previously viewed "egocentric" perspective on the layout that the cameras afforded as well as the crucial step of working out the egocentric perspective of the child between the towers (hence the second-order nature of the task). In other words, this was a task that required the integration of allocentric and egocentric information.

In light of this it is worth noting that much of the recent neuroscientific and neural-network modelling of episodic memory (e.g., Burgess et al, 2001; Lambrey et al 2012) has been focussed upon the connectivity and mutual dependency between the hippocampus (broadly responsible for allocentric spatial coding) and the medial parietal region (broadly responsible for egocentric spatial coding). Given this, it may be fruitful in this context to bridge the neuroscience and the cognitive-developmental psychology of episodic memory in the following way. The kind of early episodic memory tapped by studies such as those by Burns et al (2014) and by Newcombe et al (2014) is of a kind recruiting the hippocampus, insofar as the essential requirement in their tasks was to locate past events within an objective, allocentric framework (e.g., an action near a room landmark, in Burns et al; a toy in a certain container in Newcombe et al). Allocentric spatial facts had to be bound to temporal and semantic facts, but with no requirement to locate oneself as a viewing body within the scene. In the present study, however, while the kind of spatiotemporal coding just described did indeed have to be employed so too did a conception of the self in the past with a point-of-view. It is the parietal region rather than the hippocampus that appears to be implicated in this perspectival form of memory (Burgess et al, 2001).

The involvement of the parietal region in perspectival memory would seem to offer a neuroscientific explanation, complementing and not replacing a purely psychological one, for why we found an association between perspectival memory and theory-of-mind abilities. The medial parietal region (temporal-parietal junction) is known to be implicated in performance on theory-of-mind tasks such as false belief (e.g., Saxe and

Kanswisher, 2003). Additionally, as described in a meta-analysis by Perner and colleagues (Schurz et al, 2013) there is found to be common activation for false-belief reasoning and visual perspective taking in the left (though not the right) dorsal temporo-parietal junction.

### ***Summing Up***

We have shown that children above four-and-a-half become able to recollect and locate themselves within a past videoed scene on the basis of spatiotemporal cues. Moreover, they seem to do this not on the basis of familiarity, but by co-ordinating the point-of-view of the camera with their original point-of-view. Second, we extend and elaborate on Josef Perner's insight that the kind of episodic memory abilities emerging around four years depend upon theory-of-mind abilities, while differing from Perner on the precise nature of the latter. Finally, for reasons just outlined, it is more than mere speculation to say that maturation of the medial parietal region may underlie both the theory-of-mind abilities and the kind of perspectival episodic recollection tapped in this study.

## **APPENDIX**

A selection of justifications given by the children for choosing a particular video.

“Because that light came on first and then that light came on”

“Because they all came on together”

“Because both lights came on at the same time”

“Because the lights flashed like that”

“Because I memorized which light came on first”

“Because that one came on first”

“Because the lights came on together”

“Because it was that one then that one”

“Because it did in ours”

“Because our lights went on after each other”

“Because they flashed together like this”

“Because I remembered which light came on first”

### ACKNOWLEDGEMENTS

This research was funded by grant RG58276 from the Leverhulme Trust (UK) to J. Russell (PI), C. Russell, and N. Clayton, to whom we are most grateful. We are also grateful to Mark Haggard for advice about statistical analyses, and to the parents of the children for giving up their time.

### REFERENCES

- Babb, S. J., & Crystal, J. D. (2006). Episodic-like memory in the rat. *Current Biology*, 16(13), 1317-1321.
- Burgess, N., Becker, S., King, J. A. and O'Keefe, J. (2001) Memory for events and their spatial context: models and experiments. *Proceedings of the Royal Society of London, B*, 356, 1-11
- Burns, P., Russell, C. & Russell, J. (2014) Pre-school children's proto-episodic memory assessed by deferred imitation. *Memory* <http://dx.doi.org/10.1080/09658211.2014.963625>

- Campbell, J. (1994) *Past, space, and self*. Cambridge, MA.: MIT Press.
- Clayton, N. S. and Dickinson, A. (1998) Episodic-like memory during cache recovery by scrub jays. *Nature*, 395, 272-274.
- Drummy, A. B. and Newcombe, N. (2002) Developmental changes in source memory. *Developmental Science*, 5, 502-513
- Eacott, M. J., & Norman, G. (2004). Integrated memory for object, place, and context in rats: A possible model of episodic-like memory? *Journal of Neuroscience*, 24, 1948–1953.
- Hayne, H. & Imuta, K. (2011) Episodic memory in 3- and 4-year-old children. *Developmental Psychobiology*, 53, 317-322.
- Hoerl, C. 2001: The phenomenology of episodic recall. In T. McCormack and C. Hoerl (Eds.) *Time and Memory: Issues in Philosophy and Psychology*. Oxford: Oxford University Press.
- Iordanova, M. D., Good, M. A. & Honey, R. C. (2008) Configural learning without reinforcement: Integrated memories for correlates of what, where, and when. *Quarterly Journal of Experimental Psychology*, 61 (12), 1787-1792.
- Johnson-Laird, P.N. & Byrne, R.M.J. (1990) *Deduction*. Erlbaum, Hove & London.
- Kant, I. 1781/1998: *Critique of Pure Reason*, trans. and ed. P. Guyer and A. W. Wood. Cambridge: Cambridge University Press.
- Lambrey, S., Doeller, C., Berthoz, A. and Burgess, N. (2012) Imagining being somewhere else: Neural basis of changing perspectives in space. *Cerebral Cortex*, 22, 166-174.

McCormack, T. & Hoerl, C. (2001) Temporal concepts and self-consciousness in the development of episodic memory. In C. Moore and K. Lemmon (Eds.) *The self in time*.

Mahwah NJ: Lawrence Erlbaum.

McCormack, T., Simms, V., McGourty, J., & Beckers, T. (2013). Encouraging children to think counterfactually enhances blocking in a causal learning task. *The Quarterly Journal of Experimental Psychology*, 66, 1910-1926.

Naito, M. 2003: The relationship between theory of mind and episodic memory: Evidence for the development of autonoetic consciousness. *Journal of Experimental Child Psychology*, 85, 312–336.

Newcombe, N., & Huttenlocher, J. (1992) Children's early ability to solve perspective-taking problems. *Developmental Psychology*, 28 , 635-643.

Newcombe, N. S., Balcomb, F., Ferrara, K., Hansen, M. & Koski, J. (2014) Two rooms, two representations? Episodic-like memory in toddlers and preschoolers. *Developmental Science*, 17, 185-200.

Perner, J. (2001) Episodic memory: Essential distinctions and developmental implications. In C. Moore and K. Lemmon (Eds.) *The Self in Time*. Mahwah NJ: Lawrence Erlbaum.

Perner, J. and Howes, D. (1992) He thinks he knows: And more developmental Evidence again simulation (role taking) theory, *Mind and Language*, 7, 72-86.

Perner, J. & Ruffman, T. 1995: Episodic memory and autonoetic consciousness: Developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology*, 59, 516-548.

Perner, J., Kloo, D. & Gornik, E. 2007: Episodic memory development: Theory of mind is part of re-experiencing events. *Infant and Child Development*, 15, 25-51.

Perner, J., Kloo, D., & Stöttinger, E. 2007: Introspection and remembering. *Synthèse*, 159,253-270.

Povinelli, D.J., Landau, K.R., & Perilloux, H.K. (1996). Self-recognition in young children using delayed versus live feedback: Evidence for a developmental asynchrony. *Child Development*, 67, 1540-54.

Povinelli, D. J. , Landry, A. M. , Theall, L. A. , Clark, B. R., Castille, C. M. (1999) Development of young children's understanding that the recent past is causally bound to the present. *Developmental Psychology*, 35, 1426–1439.

Russell, J. & Hanna, R. (2012) A minimalist approach to the development of episodic memory. *Mind and Language*. 27, 29-54.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *NeuroImage*, 19, 1835-1842.

Scarf, D., Gross, L., Colombo, M (2013) To have and to hold: episodic memory in 3-and 4-year-old children, *Developmental Psychobiology*, 55 (2) 125–132.

Schurz, M., Aichhorn, M., Martin, A. and Perner, J. (2013) Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies.

Suddendorf, T. & Corballis, M. C. (1997) Mental time travel and the evolution of the human mind. *Genetic Social and General Psychology Monographs*. 123(2),133–67.



Tulving, E. (1972). Episodic and semantic memory. In: *Organisation of Memory*. E.

Tulving and W. Donaldson (Eds.), pp. 381-403. New York: Academic Press.

Tulving, E. (1985) Memory and consciousness. *Canadian Psychology*, 26, 1-12.

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human? In H. S. Terrace, & J. Metcalfe (Eds.), *The Missing Link in Cognition. Origins of Self-Reflective Consciousness* (pp. 3-56). Oxford: Oxford University Press.

Wimmer, H. and Perner, J.(1983) Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception', *Cognition* 13, 103-128.

Table 1. Percentage of correct video choice by age-group and camera position (correct choices with appropriate justification in parenthesis).

	3.5-4.5	4.5-5.5	5.5-6.5	Overall
Behind	18% (18%)	53% (41%)	77% (65%)	49% (41%)
Facing	29% (12%)	59% (41%)	47% (29%)	45% (27%)
Above	29% (12%)	41% (24%)	65% (65%)	45% (33%)
Overall	25% (14%)	51% (33%)	63% (53%)	46% (33%)

Table 2. Percentage of correct video choice by age group and light order (correct choices with appropriate justifications in parenthesis).

	3.5-4.5	4.5-5.5	5.5-6.5	Overall
L-R	30% (15%)	60% (40%)	39% (33%)	43% (30%)
Simultaneous	24% (18%)	60% (45%)	81% (69%)	55% (43%)
R-L	24% (10%)	31% (19%)	71% (59%)	41% (28%)
Overall	25% (14%)	51% (33%)	63% (53%)	46% (33%)

Table 3. This shows the distribution of reversal (bold) to simultaneous errors for children by age-group on the successive conditions.

Age group	<b>Reversal errors</b> vs. simultaneous errors
Young	<b>10</b> – 12
Middle	<b>7</b> – 8
Old	<b>2</b> - 14

Table 4. Percentage pass rates for theory of mind, Modus Tollens, perspective-taking and know/guess.

Age Group ( <i>n</i> = 51 per group)	ToM	Modus Tollens	Perspective-Taking		Know/Guess	
			0-2	3-4	0-1	1.5-2
3.5-4.5	29	86	61	39	84	16
4.5-5.5	51	86	55	45	64	36
5.5-6.5	84	98	12	88	8	92
Overall	54	90	47	53	53	47

Table 5. Raw correlations (above diagonal) and partial correlations controlling for age and BPVS score (below diagonal).

	1.	2.	3.	4.	5.	6.	7.
1. Age	-	.59**	.35**	.25**	.64**	.38**	.21*
2. BPVS		-	.43**	.18*	.43**	.25**	.3**
3. Jungle Task			-	.18*	.26**	.21**	.15
4. ToM			.18*	-	.43**	.18*	-.19*
5. Know/guess			.03	.16*	-	.36**	.28**
6. Perspective			.09	.27**	.17*	-	.22**
7. Modus T.			.02	.09	.19*	.14	-

Note: \* = significant at the .05 level; \*\* = significant at the .01 level.

Table 6. Final logistic regression model with performance on the jungle task (pass, fail) as the dependent variable and age (months), BPVS (raw score) and second-order theory of mind (pass, fail) the predictor variables.

	<i>B</i>	<i>SE β</i>	<i>Wald's χ<sup>2</sup></i>	<i>df</i>	<i>p</i>	<i>e<sup>B</sup></i> ( <i>odds ratio</i> )
Constant	-5.14	1.19	18.68	1	.00	.01
Age	.02	.02	1.18	1	.28	1.02
BPVS	.04	.02	6.24	1	.01	1.04
Theory of Mind	1.1	.48	5.25	1	.02	2.98

Figure 1. Layout of the “jungle” room

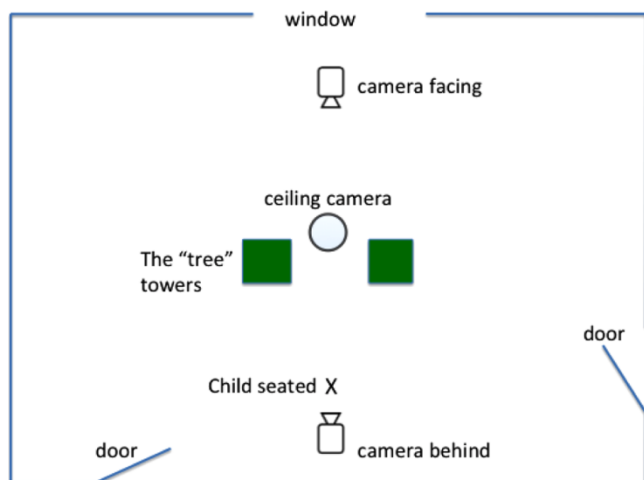




Figure 2. A view of the “jungle” room

